Tilburg University

Nonparametric Nonlinear Cotrending Analysis, with an Application to Interest and Inflation in the U.S

Bierens, H.J.

Publication date:
1996

Link to publication in Tilburg University Research Portal

Citation for published version (APA):
Bierens, H. J. (1996). *Nonparametric Nonlinear Cotrending Analysis, with an Application to Interest and Inflation in the U.S*. (CentER Discussion Paper; Vol. 1996-62). Econometrics.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 05. May. 2023

NONPARAMETRIC NONLINEAR CO-TRENDING ANALYSIS,

WITH AN APPLICATION TO INTEREST AND INFLATION IN THE U.S.

Herman J. Bierens[1]

Southern Methodist University, Dallas, U.S.A.

& Tilburg University, the Netherlands

Abstract:

Given the assumption that the components of a vector time series are stationary about nonlinear deterministic time trends, nonlinear co-trending is the phenomenon that one or more linear combinations of the time series are stationary about a linear trend, hence the series have common nonlinear deterministic time trends. In this paper we shall develop nonparametric tests for nonlinear co-trending. The tests are based on generalized eigenvalues, where the two matrices involved are constructed nonparametrically on the basis of partial sums. We apply this approach to the federal funds rate and the CPI inflation rate in the U.S., using monthly data. It appears that these series are nonlinearly co-trended, where the nonlinear trend in the inflation rate is positively related to the nonlinear trend in the interest rate. This positive relation between interest and inflation is known in the literature as the *price puzzle*. Thus, our result suggests that the price puzzle is due to a common nonlinear time trend in the series involved.

Current version: augustus 7, 1996

[1] Correspondence address until December 1996: Department of Economics, Southern Methodist University, Dallas, TX 75275, U.S.A. E-Mail: hbierens@mail.smu.edu. From January 1997 onwards: Department of Economics, Penn State University, University Park, PA 16802, U.S.A. A previous version of this paper was presented at York University, Canada, Texas Econometrics Camp 1996, and Penn State University.

1. INTRODUCTION

The aim of this paper is twofold. Our first aim is to develop nonparametric tests for nonlinear co-trending of macroeconomic time series. Given the assumption that the components of a vector time series are stationary about nonlinear deterministic time trends, nonlinear co-trending is the phenomenon that one or more linear combinations of the time series are stationary about a linear trend, hence the series have common nonlinear deterministic time trends. Second, we want to investigate the nature of the relation between the federal funds rate and the CPI inflation rate in the U.S., in particular whether this relation is due to a common nonlinear deterministic time trend. Since the 1950's these two macroeconomic time series show a remarkable similarity, known as the *price puzzle*. See Bernanke and Blinder (1992), Christiano and Eichenbaum (1992), Christiano, Eichenbaum and Evans (1992, 1995), Eichenbaum (1994), Sims (1995), and Balke and Emery (1994, 1995).

The kind of nonlinear trend stationarity we consider in this paper is $z_t = \beta_0 + \beta_1 t + f(t) + u_t$, where z_t is a k-variate time series process, u_t is a k-variate zero-mean stationary process, and $f(t)$ is a deterministic k-variate nonlinear trend function represening structural change. Nonlinear co-trending is then the phenomenon that there exists a non-zero vector θ such that $\theta^T f(t) = 0$.

The motivation for considering nonlinear co-trending is threefold. First, there is now empirical evidence that some long macro-economic time series such as those in the Nelson-Plosser (1982) data set that were initially perceived as unit root processes are probably more in accordance with a nonlinear trend stationary hypothesis. See for example Perron (1988, 1989, 1990) who tested the unit root hypothesis for the Nelson-Plosser series against trend stationarity with a trend break (which is a special case of nonlinear trend stationarity), and Bierens (1996a) who tested the unit root hypothesis for the price level and interest rate series in the extended Nelson-Plosser data set (extended by Schotman and Van Dijk (1991) to 1988) against a smooth nonlinear trend stationarity hypothesis. However, despite the somewhat reluctant conclusion of Bierens (1996a) that the log of the annual CPI over the period 1860-1988 is probably a nonlinear trend stationary process rather than a unit root process, it appears that for monthly post-war time series the log of the CPI looks more like a unit root process with time varying drift, hence the CPI inflation rate is then a nonlinear trend stationary process.

The second motivation is that quite a few macroeconomic time series that are not unit root processes still behave like cointegrated processes in that the series move together over time in a similar way. But cointegration is only possible for unit root processes, so something else is going on. A possible explanation is that these series have common nonlinear deterministic time trends.

The third motivation is that the (linear trend) stationarity hypothesis as well as the unit root (with constant drift) hypothesis for macroeconomic time series imply that the structure of the economy (i.e. the parameters of the underlying data-generating process) does not change over time. This is quite implausible, in particular for long macro-economic time series such as the Nelson-Plosser data spanning a century or more.

The plan of the paper is as follows. In section 2 we summarize the procedure for testing the number of co-trending vectors and the ideas behind it. In particular, we show how to construct nonparametrically two matrices $\hat{M}_1$ and $\hat{M}_2$ such that their generalized eigenvalues can be used to test for nonlinear co-trending. In section 3 we discuss some properties of partial sums of the nonlinear trend function, and in section 4 we derive the asymptotic properties of the matrices $\hat{M}_1$ and $\hat{M}_2$. In section 5 we derive the actual tests for the number of co-trending vectors on the basis of the generalized eigenvalues of the matrices $\hat{M}_1$ and $\hat{M}_2$, and the asymptotic null distributions of the tests. Also, we propose a test of linear restrictions on the co-trending vectors. In section 6 we propose consistent estimators of the co-trending vectors. In section 7 we show what happens if our tests are applied to a cointegrated unit root process rather than a nonlinear trend stationary process. In section 8 we apply our approach to monthly time series of the federal funds rate and the CPI inflation rate in the U.S. Most of the proofs are given in the Appendex.

2. INTRODUCTION TO NONLINEAR CO-TRENDING ANALYSIS

Consider a k-variate time series process $z_t = g(t) + u_t$, where $g(t) = E(z_t)$ is a nonlinear trend function and u_t is a zero mean stationary process. In this paper we shall design a test of the null hypothesis that there exists a nonzero k-vector θ such that $\theta^T g(t)$ is linear in t; in other words, we test the null hypothesis that the time series z_t is nonlinear co-trended. First, we shall design a test of the null hypothesis $\mathfrak{C}(1)$ that the space of all such co-trending vectors θ has dimension 1, against the alternative $\mathfrak{C}(0)$ that this dimension is zero, i.e., we test $\mathfrak{C}(1)$ against the alternative hypothesis that the only vector θ for which $\theta^T g(t)$ is linear in t is the zero vector. Subsequently, we shall extend this test to testing $\mathfrak{C}(r)$ against $\mathfrak{C}(s)$ with $0 \leq s < r$, for $r = 1,..,k$.

For $t = 1,....,n$ we can always write $g(t) = \beta_{0,n} + \beta_{1,n}t + f_n(t)$, where f_n(t) is such that

$$(1) \qquad \sum_{t=1}^{n} f_n(t) \quad 0, \quad \sum_{t=1}^{n} t\, f_n(t) \quad 0.$$

Thus, $\theta^T g(t)$ is linear in t if and only if for all $n \geq 1$ and $t = 1,..,n$, $\theta^T f_n(t) = 0$, and consequently, denoting

$$(2) \qquad {}_{l}(x) \quad (1/n)\sum_{t=1}^{[nx]} f_n(t) \;\; if \;\; x \in [n^{-1},1], \quad F_n(x) \quad 0 \;\; if \;\; x \in [0,n^{-1}$$

z_t is nonlinear co-trended if $\theta^T F_n(x) = 0$ for all x in [0,1] and $n \geq 1$.

Note that (1) and (2) imply

$$(3) \qquad F_n(0) \quad F_n(1) \quad 0, \quad \int F_n(x)dx \quad 0,$$

where the latter result follows from Lemma 9.6.3. in Bierens (1994, p.200). The integral in (3) is taken over the unit interval, as will be in the sequel unless otherwise indicated.

We shall establish conditions such that $F_n(x) \rightarrow F(x)$ in L^2 (w.r.t. the Lebesgue measure on [0,1]). Then under the nonlinear co-trending hypothesis $\mathfrak{C}$ (1) the vector θ is the eigenvector of the matrix

$$M_1 \quad \int F(x)F(x)^T dx$$

corresponding to a zero eigenvalue.

We propose to estimate the matrix M_1 by

$$\hat{M}_1 \quad (1/n)\sum_{t=1}^{n} \hat{F}(t/n)\hat{F}(t/n)^T$$

where

$$\hat{F}(x) \quad (1/n)\sum_{t=1}^{[nx]} (z_t \quad \hat{\beta}_0 \quad \hat{\beta}_1 t) \;\; if \;\; x \in (n^{-1},1],$$
$$\hat{F}(x) \quad 0 \;\; if \;\; x \in [0,n^{-1}],$$

with $\hat{\beta}_0$ and $\hat{\beta}_1$ the OLS estimates of the vectors of intercepts and slope parameters in the regression of z_t on time t for $t = 1,...,n$. Note that, since $\hat{F}(x)$ is a step function and $\hat{F}(0) = \hat{F}(1) = 0$,

$$\hat{M}_1 \quad \int \hat{F}(x)\hat{F}(x)^T dx.$$

It will be shown that under the nonlinear co-trending hypothesis $\mathfrak{C}$ (1) and the assumption that u_t has the Wold representation

(4) $$u_t \quad C(L)\varepsilon_t \quad \sum_{j=0}^{\infty} C_j\varepsilon_{t-j},\ where\ \varepsilon_t \sim i.i.d.\ (0,I_k),$$

$\hat{M}_1 \rightarrow M_1$ in probability, and in particular that $n\theta^T\hat{M}_1\theta$ converges in distribution to a functional of a standard Wiener process, times $\theta^T C(1)C(1)^T\theta$. The latter is a kind of nuisance parameter which we want to get rid of by using a Newey-West (1987) type estimator for $\theta^T C(1)C(1)^T\theta$, along the lines in Bierens (1994, p.197), as follows. Consider the matrix

(5) $$(1/n)\sum_{t=m}^{n}\left((1/m)\sum_{j=0}^{m-1}\left(z_{t-j}\ \hat{\beta}_0\ \hat{\beta}_1(t\ j)\right)\right)\left((1/m)\sum_{j=0}^{m-1}\left(z_{t-j}\ \hat{\beta}_0\ \hat{\beta}_1(t\ j)\right.\right.$$

where $m = [n^{\alpha}]$ with $0 < \alpha < 1$. It will be shown that $\hat{M}_2 \rightarrow M_2 = \int F'(x)F'(x)^T dx$, and under $\mathfrak{C}(1)$, $n^{\alpha}\theta^T\hat{M}_2\theta \rightarrow \theta^T C(1)C(1)^T\theta$ in prob. A test of $\mathfrak{C}(1)$ against $\mathfrak{C}(0)$ can now be based on the minimum solution $\hat{\lambda}_1$, say, of the generalized eigenvalue problem $\det[\hat{M}_1 - \lambda\hat{M}_2] = 0$; i.e., the test statistic involved is $n^{1-\alpha}\hat{\lambda}_1$. The reason for using this generalized eigenvalue approach is that then asymptotically $\theta^T C(1)C(1)^T\theta$ will cancel out. We shall extend this test to the case of multiple nonlinear co-trending, and to testing linear restrictions on the co-trending vectors θ. The asymptotic power of these tests depends on the choice of α: the smaller α, the higher the asymptotic power. However, the rate of convergence of $n^{\alpha}\theta^T\hat{M}_2\theta$ to $\theta^T C(1)C(1)^T\theta$ is optimal for $\alpha \quad 1/2$, hence too small an α may cause size distortion. In the empirical application we shall therefore choose $\alpha \quad 1/2$.

Also, we show that under the hypothesis $\mathfrak{C}(r)$ with $r \geq 1$ the eigenvectors of the matrix $\hat{M}_1$ corresponding to the r smallest eigenvalues are $\sqrt{n}$-consistent estimators of the co-trending vectors θ.

Similar to the cointegration tests of Johansen (1988,1991,1994), Johansen and Juselius (1990), and Bierens (1996b), the following result of Anderson, Brons and Jensen (1983) plays a key role in the derivation of the asymptotic null distribution:

If a pair of square random matrices (P_n, Q_n) *converges in distribution to* (P,Q), *where Q is a.s. nonsingular, then the ordered solutions of the generalized eigenvalue problem* $\det[P_n - \lambda Q_n] = 0$ *converge in distribution to the ordered solutions of the generalized eigenvalue problem* $\det[P - \lambda Q] = 0$.

This result cannot be applied directly to the matrices $\tilde{M}_1$ and $\tilde{M}_2$, because under $\mathfrak{C}(1)$ both matrices converge in distribution to singular matrices, but it is applicable to their inverses after some rescaling. This is the reason why in the discussion on the asymptotic properties of the matrices $\tilde{M}_1$ and $\tilde{M}_2$ we will focus on their inverses.

3. NONLINEAR DETERMINISTIC TRENDS AND THEIR PARTIAL SUMS

Without loss of generality we may write the nonlinear trend function $f_n(t)$ as a series expansion of orthogonal Chebishev polynomials (cf. Hamming 1973):

$$f_n(t) = \sum_{j=1}^{n-1} \gamma_{j,n}\sqrt{2}\cos\left(j\pi\left(t - \tfrac{1}{2}\right)/n\right), \quad t = 1,\ldots\ldots,n,$$

where

$$\gamma_{j,n} = (1/n)\sum_{t=1}^{n} f_n(t)\sqrt{2}\cos\left(j\pi\left(t - \tfrac{1}{2}\right)/n\right).$$

Note that for $j, j_1, j_2 = 1,\ldots n\text{-}1$,

$$(1/n)\sum_{t=1}^{n} \sqrt{2}\cos\left(j\pi\left(t - \tfrac{1}{2}\right)/n\right) = 0$$

and

$$(6) \qquad (1/n)\sum_{t=1}^{n} 2\cos\left(j_1\pi\left(t - \tfrac{1}{2}\right)/n\right)\cos\left(j_2\pi\left(t - \tfrac{1}{2}\right)/n\right) = I(j_1 = j_2),$$

where $I(.)$ is the indicator function. Now let

$$(7) \qquad \phi_n(x) = f_n(nx) = \sum_{j=1}^{n-1} \sqrt{2}\gamma_{j,n}\cos\left(j\pi\left(x - \tfrac{1}{2n}\right)\right), \quad x \in [0,1],$$

and assume:

ASSUMPTION 1. $\phi_n(x)$ *converges to* $\phi(x)$ *in* L^2, *i.e.,* $\int \|\phi_n(x) - \phi(x)\|^2 dx = o(1)$, *where* ϕ *is* *a twice differentiable vector function of the form*

$$\phi(x) \quad \sum_{j=1}^{\infty} \gamma_j \sqrt{2} \cos(j\pi x) .$$

Moreover, the vectors $\phi_n^{//}$ *and* $\phi^{//}$ *of second derivatives of the components of* ϕ_n *and* ϕ*, respectively, satisfy*

$$(1/n)\sum_{t=1}^{n} \|\phi_n^{//}(t/n)\|^2 \quad O(1) , \quad \int \|\phi^{//}(x)\|^2 dx < \infty . \tag{8}$$

Furthermore, under the hypothesis $\mathfrak{C}(r)$*,*

$$\mathrm{nk}\left[\int \left(\int_0^x \phi(y)dy\right)\left(\int_0^x \phi(y)dy\right)^T dx\right] \quad \mathrm{rank}\left[\int \phi(x)\phi(x)^T dx\right] \quad k \tag{9}$$

The conditions in (8) are smoothness conditions, and condition (9) is a regularity condition which ensures that under the hypothesis $\mathfrak{C}(r)$ the rank of the matrix M_1 and the probability limit of the matrix $\tilde{M}_2$ is k-r.

The following two lemmas now provide some preliminary convergence results for the vector function $F_n(.)$:

LEMMA 1. *Under Assumption 1,* $\int \|F_n(x) \quad F(x)\|^2 dx \quad o(1)$ *, where*

$$F(x) \quad \int_0^x \phi(y)dy \quad \frac{\sqrt{2}}{\pi}\sum_{j=1}^{\infty} \frac{\gamma_j}{j} \sin(j\pi x) , \tag{10}$$

Proof: Appendix.

Note that Assumption 1 implies that $F(x)$ is differentiable:

$$F^{/}(x) \quad \sqrt{2}\sum_{j=1}^{\infty} \gamma_j \cos(j\pi x) \quad \phi(x) .$$

LEMMA 2. *Let* δ_n *be a sequence of scalars such that* $\delta_n \to 0$*,* $\delta_n n \to \infty$*. Then under Assumption 1,* $\delta_n^{-1}\left(F_n(x \ \delta_n) \quad F_n(x)\right) \quad \phi_n(x) = O(\delta_n) \quad O\left(\delta_n^{-1} n^{-1}\right) = o(1)$*, uniformly on* $[0,1]$.

Proof: Appendix.

Note that Lemma 2 implies that $\int \| \delta_n^{-1}\left(F_n(x+\delta_n) - F_n(x)\right) - F'(x)\|^2 dx = o_p(1)$.

4. THE ASYMPTOTIC PROPERTIES OF THE MATRICES $\hat{M}_1$ AND $\hat{M}_2$

Given the Wold representation (4), we can always write:

$$
\begin{aligned}
u_t &= C(L)\varepsilon_t = C(1)\varepsilon_t + (1-L)\frac{C(L)-C(1)}{1-L}\varepsilon_t \\
&= C(1)\varepsilon_t + (1-L)D(L)\varepsilon_t = C(1)\varepsilon_t + v_t - v_{t-1},
\end{aligned}
\tag{11}
$$

say, where $D(L) = [C(L)\text{-}C(1)]/(1\text{-}L)$ and $v_t = D(L)\varepsilon_t$. A sufficient condition for the stationarity of u_t and v_t is that:

ASSUMPTION 2. *The process u_t has the Wold representation* (4), *where the ε_t's are i.i.d. $N_k(0,I)$, and $C(L) = C_1(L)^{-1}C_2(L)$, where $C_1(L)$ and $C_2(L)$ are matrix-valued finite-order lag polynomials, such that all the roots of* $\det(C_1(L))$ *lie outside the unit circle.*

Cf. Engle (1987). This assumption is more restrictive than necessary, but it will keep the argument below transparent, and focussed on the main issues. See Phillips and Solo (1992) for weaker conditions in the case of linear processes. Also, we could assume instead of Assumption 2 that u_t is stationary and ergodic, so that we still can write $u_t = C(1)\varepsilon_t + v_t - v_{t-1}$, where now ε_t is a martingale difference process with unit variance matrix and v_t is a stationary process. Cf. Hall and Heyde (1980, p.136). Note that Assumption 2 does not restrict the lag polynomial $C_2(L)$. However, we do need the additional condition that

ASSUMPTION 3. *The matrix $C(1)$ is nonsingular.*

This separation of conditions will prove convenient when we compare nonlinear co-trending with cointegration, in section 7.

It follows now from the functional central limit theorem that, with

$$U_n(x) \quad (1/\sqrt{n})\sum_{t=1}^{[nx]} u_t \;\; if \;\; x \in [n^{-1},1], \quad U_n(x) \quad 0 \;\; if \;\; x \in [0,n^{-1})$$

we have

$$U_n(x) \Rightarrow C(1)W_k(x), \tag{12}$$

where W_k is a k-variate standard Wiener process. Cf. Billingsley (1968). Moreover, using (12) and Lemma 9.6.3 in Bierens (1994, p.200), it is a standard exercise to show that

$$\begin{aligned} F_n(x)) \Rightarrow C(1)\{W_k(x) \quad xW_k(1) \quad 3(x^2 \quad x)[2\int W_k(y)dy \\ C(1)\overline{W}_k(x), \end{aligned} \tag{13}$$

say. Then:

LEMMA 3. *Under Assumptions 1-3,*

$$\hat{M}_1 \quad M_1 \quad o_p(1), \;\; where \;\; M_1 \quad \int F(x)F(x)^T dx, \tag{14}$$

and

$$\hat{M}_2 \quad M_2 \quad o(1), \quad where \;\; M_2 \quad \int F^{\prime}(x)F^{\prime}(x)^T dx. \tag{15}$$

Proof: Appendix.

Now let Q be the orthogonal matrix of eigenvectors of M_1 corresponding to the ordered eigenvalues $\lambda_1 \le \lambda_2 \le .... \le \lambda_k$. Note that under the nonlinear co-trending hypothesis $\mathfrak{C}$ (1), λ_1 = 0, and the first column of Q is θ: $Q = (\theta, Q_*)$, say. Denoting $\Lambda_* = \mathrm{diag}(\lambda_2,..,\lambda_k)$, it follows that

LEMMA 4. *Let Assumptions 1-3 and the hypothesis* $\mathfrak{C}$ (1) *hold. For every subsequence* $m = o(n)$ *of natural numbers we have*:

$$\begin{pmatrix} 1/\sqrt{n} & 0^T \\ 0 & \sqrt{m/n}I_{k-1} \end{pmatrix} Q^T \hat{M}_1^{-1} Q \begin{pmatrix} 1/\sqrt{n} & 0^T \\ 0 & \sqrt{m/n}I_{k-1} \end{pmatrix} \rightarrow \begin{pmatrix} \tilde{\mu}^{-1} & 0^T \\ 0 & O \end{pmatrix},$$

in distribution, where

$$\tilde{\mu} \quad \theta^T C(1)\int \overline{W}_k(x)\overline{W}_k(x)^T dx C(1)^T\theta \quad \theta^T C(1)\int \overline{W}_k(x)F(x)^T dx Q_*\Lambda_*^{-1}Q_*^{\,T}\int F(y)\overline{W}_k(y)^T dy C(1)^T\theta$$

$$\sim \; \theta^T C(1)C(1)^T\theta\left(\int\left(\overline{W}_1(x)\right)^2 dx \quad \int \overline{W}_1(x)F(x)^T dx Q_*\Lambda_*^{-1}Q_*^{\,T}\int F(y)\overline{W}_1(y)dy\right)$$

Proof: Appendix.

LEMMA 5. *Let* $\hat{M}_2$ *be defined by* (5) *with* $m = [n^{\alpha}]$ *for some* $\alpha \in (0,1)$. *Under Assumptions 1-3 and hypothesis* $\mathfrak{C}(1)$,

$$\begin{pmatrix} 1/\sqrt{m} & 0^T \\ 0 & I_{k-1}\end{pmatrix} Q^{\,T}\hat{M}_2^{-1}Q\begin{pmatrix} 1/\sqrt{m} & 0^T \\ 0 & I_{k-1}\end{pmatrix} \to \begin{pmatrix} \left(\theta^T C(1)C(1)^T\theta\right)^{-1} & 0^T \\ 0 & \left(Q_*^{\,T}M_2Q_*\right)^{-1}\end{pmatrix}$$

in probability.

Proof: Appendix.

5. THE TESTS

5.1. The generalized eigenvalue problem

Now take m in Lemma 4 the same as in Lemma 5. Then the result of Lemma 4 reads:

$$\begin{pmatrix} 1/\sqrt{m} & 0^T \\ 0 & I_{k-1}\end{pmatrix} Q^{\,T}\left(\frac{m}{n}\hat{M}_1^{-1}\right)Q\begin{pmatrix} 1/\sqrt{m} & 0^T \\ 0 & I_{k-1}\end{pmatrix} \to \begin{pmatrix} \tilde{\mu}^{-1} & 0^T \\ 0 & O\end{pmatrix}$$

in distr. Hence it follows from Lemma 5 that the ordered solutions $\hat{\lambda}_1 \le \hat{\lambda}_2 \le .... \le \hat{\lambda}_k$ of the generalized eigenvalue problem

$$\det\left(\hat{M}_1 \quad \lambda\,\hat{M}_2\right) \quad 0 \tag{16}$$

are related to the decreasingly ordered solutions $\tilde{\lambda}_1 \ge \tilde{\lambda}_2 \ge .... \ge \tilde{\lambda}_k$ of the generalized eigenvalue problem

$$\det\left[\begin{pmatrix} 1/\sqrt{m} & 0^T \\ 0 & I_{k-1}\end{pmatrix} Q^{\,T}\frac{m}{n}\hat{M}_1^{-1}Q\begin{pmatrix} 1/\sqrt{m} & 0^T \\ 0 & I_{k-1}\end{pmatrix} \quad \lambda\begin{pmatrix} 1/\sqrt{m} & 0^T \\ 0 & I_{k-1}\end{pmatrix} Q^{\,T}\hat{M}_2^{-1}Q\begin{pmatrix} 1/\sqrt{m} & 0^T \\ 0 & I_{k-1}\end{pmatrix}\right] \quad 0 \tag{17}$$

by the equality:

$$\frac{n}{m}\hat{\lambda}_j \quad \tilde{\lambda}_j^{-1},\ j \quad 1,\dots,k.$$

Moreover, it follows from Lemmas 4-5 and Anderson, Brons and Jensen (1983) that the ordered solutions of generalized eigenvalue problem (17) converge in distribution to the ordered solutions of the generalized eigenvalue problem

$$\det\left[\begin{pmatrix} \tilde{\mu}^{-1} & 0^T \\ 0 & O \end{pmatrix} \quad \lambda \begin{pmatrix} \left(\theta^T C(1)C(1)^T\theta\right)^{-1} & 0^T \\ 0 & \Lambda_*^{-1} \end{pmatrix}\right] \quad 0.$$

Similarly, comparing the solutions of generalized eigenvalue problem

$$\det\left[\begin{pmatrix} \sqrt{m} & 0^T \\ 0 & I_{k-1} \end{pmatrix} Q^T \hat{M}_1 Q \begin{pmatrix} \sqrt{m} & 0^T \\ 0 & I_{k-1} \end{pmatrix}\right.$$

$$\left.\lambda \begin{pmatrix} \sqrt{m} & 0^T \\ 0 & I_{k-1} \end{pmatrix} Q^T \hat{M}_2 Q \begin{pmatrix} \sqrt{m} & 0^T \\ 0 & I_{k-1} \end{pmatrix}\right] \quad 0$$

with the solutions of eigenvalue problem (16), it follows easily that under Assumptions 1-3 and the hypothesis $\mathfrak{H}(1)$, the vector $(\hat{\lambda}_2,..,\hat{\lambda}_k)$ converges in probability to the vector of ordered solutions of the generalized eigenvalue problem $\det[\Lambda_* \quad \lambda Q_*^T M_1 Q_*] \quad 0$. Thus we have

THEOREM 1: *Let $\hat{M}_2$ be defined as in Lemma 4. Under Assumptions 1-3 and hypothesis* $\mathfrak{H}(1)$, *the minimum solution $\hat{\lambda}_1$ of the generalized eigenvalue problem* (16) *satisfies*

$$(18) \qquad {}^{1-\alpha}\hat{\lambda}_1 \rightarrow \int\left(\overline{W}_1(x)\right)^2 dx \quad \int \overline{W}_1(x)F(x)^T dx Q_* \Lambda_*^{-1} Q_*^T \int F(x)\overline{W}_1(x)^T d$$

in distribution, where $\overline{W}_1$ is defined in (13), *whereas under the hypothesis* $\mathfrak{H}(0)$, $\hat{\lambda}_1$ *converges in probability to a positive constant.*

We recall that the power of the test depends on the choice of $\alpha \in (0,1)$: the smaller α the higher the asymptotic power. However, the value α = 1/2, hence $m = [\sqrt{n}]$ is optimal for the rate of convergence of $m\theta^T\hat{M}_2\theta$ to $\theta^T C(1)C(1)^T\theta$, so that we may expect that choosing α too far away from 1/2 will cause size distortion in samples of moderate size.

5.2. An upperbound of the asymptotic size of the test

A practical problem with this test is that the null distribution is case-dependent: it depends on F. A simple way to get around this problem is to base upperbounds of the asymptotic critical values of the test on the random variable

$$\int\left(\bar{W}_1(x)\right)^2 dx \tag{19}$$

only. A motivation for this choice is that under $\mathfrak{C}$ (1), the function $F(x)$ may be considered as a parameter in a space Ξ_1, where

DEFINITION 1: Ξ_r *is the space of k-dimensional continuous functions F on* [0,1] *for which* $F(0) = F(1) = 0$, $\int F(x)dx = 0$, $\int F(x)^T F(x)dx < \infty$, *and* $\text{rank}[\int F(x)F(x)^T dx] = k\text{-}r$,

and that therefore the null hypothesis of co-trending is a composite hypothesis. Then:

LEMMA 6: *There exists a sequence F_m in Ξ_1 such that*

$$\underset{m\to\infty}{\text{plim}}\int \bar{W}_1(x)F_m(x)dx \quad 0.$$

hence (19) *is the supremum of the right hand side of* (18) *over all F in* Ξ_1.

Proof: Appendix.

Thus we have:

THEOREM 2: *Under the conditions of Theorem 1,*

$$\sup_{F\in\Xi_1}\ \lim_{n\to\infty} P\left(n^{1-\alpha}\hat{\lambda}_1 \ \geq\ K\right) \quad P\left(\int\left(\bar{W}_1(x)\right)^2 dx \ \geq\ K\right),$$

for every $K > 0$.

5.3. Multiple nonlinear co-trending

The extension of the above test to testing $\mathfrak{C}$ (r) against $\mathfrak{C}$ (s) with $s < r$ is straightforward. Let $\Lambda_{k\text{-}r}$ be the diagonal matrix of the k-r largest eigenvalues of the matrix M_1, let $Q_{k\text{-}r}$ be the matrix of corresponding orthonormal eigenvectors, let $\bar{\lambda}_r$ be the maximum eigenvalue of the matrix

$$\int \overline{W}_r(x)\overline{W}_r(x)^T dx \quad \int \overline{W}_r(x)F(x)^T dx\, Q_{k-r}\Lambda_{k-r}^{-1}Q_{k-r}^T \int F(y)\overline{W}_r(y)^T dy,$$

and let $\overline{\lambda}_r^*$ be the maximum eigenvalue of the matrix

$$\int \overline{W}_r(x)\overline{W}_r(x)^T dx.$$

Then it is easy to show:

THEOREM 3. *Under Assumptions 1-3 and hypothesis* $\mathfrak{C}(r)$, $n^{1-\alpha}\hat{\lambda}_r \rightarrow \overline{\lambda}_r$ *in distr., and for every* $K > 0$, $\sup_{F\in\Xi_r} \lim_{n\rightarrow\infty} P\left[n^{1-\alpha}\hat{\lambda}_r \geq K\right] \quad P\left[\overline{\lambda}_r^* \geq K\right]$, *whereas under hypothesis* $\mathfrak{C}(s)$ *with* $s < r$, $\hat{\lambda}_r$ *converges in probability to a positive constant.*

The 80%, 90% and 95% quantiles of the distribution of the random variable $\overline{\lambda}_r^*$ for $r = 1,..,5$ are given in Table 1. These quantiles are calculated by Monte Carlo simulation, on the basis of 10,000 replications of samples of size $n = 500$ from the $N_r(0,I_r)$ distribution.

Table 1: *Values of K for which* $P(\overline{\lambda}_r^* \leq K) \quad p$

p:	0.80	0.90	0.95
r:		K	
1	0.091103	0.119616	0.150989
2	0.134492	0.169183	0.202642
3	0.173114	0.214069	0.252212
4	0.205922	0.251317	0.294746
5	0.236006	0.282870	0.330943

5.4. Testing linear restrictions on the co-trending vectors

Once we have established the number r of linear independent co-trending vectors θ, we may wish to test the null hypothesis that the columns of a given $k \times s$ matrix H with $1 \leq s \leq r$ span a subspace of the space of co-trending vectors, similar to testing linear restrictions on the cointegrating vectors by Johansen's (1988,1991,1994) likelihood ratio approach. It is straightforward to verify from the proofs of lemmas 4 and 5, in particular the equations (A.9) and (A.10), that under this hypothesis

$$nH^T\hat{M}_1H \to H^TC(1)\int\overline{W}_k(x)\overline{W}_k(x)^TdxC(1)^TH$$
$$\sim \left(H^TC(1)C(1)^TH\right)^{1/2}\left(\int\overline{W}_s(x)\overline{W}_s(x)^Tdx\right)\left(H^TC(1)C(1)^TH\right)^{1/2}$$

in distr., and with $m = [n^\alpha]$, $0 < \alpha < 1$,

$$n^\alpha H^T\hat{M}_2H \to H^TC(1)C(1)^TH$$

in prob. Now let $\tilde{\lambda}_s$ be the maximum solution of the generalized eigenvalue problem

$$\det[H^T\hat{M}_1H - \lambda H^T\hat{M}_2H] = 0.$$

Then it is not hard to verify:

THEOREM 4. *Let Assumptions 1-3 hold, and let z_t be nonlinear co-trended with r co-trending vectors. Let H be a given $k \times s$ matrix with $1 \leq s \leq r$. If the columns of H span a subspace of the space of co-trending vectors then $n^{1-\alpha}\tilde{\lambda}_s$ converges in distribution to the maximum eigenvalue $\overline{\lambda}_s^*$ of the matrix $\int\overline{W}_s(x)\overline{W}_s(x)^Tdx$, whereas otherwise $\tilde{\lambda}_s$ converges in probability to a positive constant.*

5.5 Detrending or not?

All our results so far are based on detrended data. But there are situations where there is no linear trend in the data, for example when we use differenced time series. Then taking out a constant mean, by subtracting from z_t its sample mean $\overline{z}$, will suffice. It is easy to verify that all our results carry over, provided that we replace the process $\overline{W}_k(x)$ defined in (13) by a k-variate standard Brownian bridge $W_k^o(x) = W_k(x) - xW_k(1)$. Moreover, instead of the critical values in Table 1 we should use the ones in Table 2 below, where λ_r^o is the maximum eigenvalue of the matrix

$\int W_r^o(x) W_r^o(x)^T dx$.

Table 2: Values of K for which $P(\lambda_r^o \le K)$ $\;\;p$

p:	0.80	0.90	0.95
r:		K	
1	0.2451126	0.3518246	0.4657737
2	0.3993106	0.5356136	0.6742039
3	0.5413243	0.7036614	0.8603746
4	0.6778114	0.8618191	1.0345377
5	0.8170006	1.0141629	1.2194813

6. CONSISTENT ESTIMATION OF THE CO-TRENDING VECTORS

Given the hypothesis $\mathfrak{C}(1)$, there are three candidates for the estimator $\hat\theta$, say, of the co-trending vector θ, namely the eigenvector corresponding to the minimum eigenvalue of the matrix $\hat M_1$, the minimum eigenvalue of the matrix $\hat M_2$, or the minimum generalized eigenvalue of $\hat M_1$ w.r.t. $\hat M_2$. Let us first derive the limiting distribution of $\hat\theta$ in the case $\hat M_1$. Let $\hat\xi$ be the eigenvector corresponding to the minimum eigenvalue of $Q^T \hat M_1 Q$, normalized such that $\hat\xi^T \quad (1, \hat\xi_*^T)$, and let $\tilde\theta \quad Q\hat\xi$. Note that $\tilde\theta \quad Q\hat\xi$ is an eigenvector of $\hat M_1$ corresponding to the minimum eigenvalue. Then $Q_*^T \hat M_1 \theta \quad Q_*^T \hat M_1 Q_* \hat\xi_* \quad \hat\lambda_1 \hat\xi_*$, hence it follows from the proof of Lemma 4, in particular equation (A.9), that

$$\sqrt{n}\,\hat\xi_* \quad [Q_*^T \hat M_1 Q_* \quad \hat\lambda_1 I_{k-1}]^{-1} \sqrt{n}\, Q_*^T \hat M_1 \theta \rightarrow \quad \Lambda_*^{-1} Q_*^T \int F(x) \overline{W}_k(x)^T dx C(1)^T \theta$$

in distr., and consequently,

$$\sqrt{n}(\tilde\theta \quad \theta) \quad Q_* \sqrt{n}\hat\xi_* \rightarrow \quad Q_* \Lambda_*^{-1} Q_*^T \int F(x) \overline{W}_k(x)^T dx C(1)^T \theta \tag{20}$$

in distr. Moreover, by the mean value theorem,

$$\|\tilde{\theta}\| \quad 1 \quad \sqrt{1 \;\hat{\xi}_*^T Q_* Q_*^T \hat{\xi}_*} \quad 1 \quad O_p(1/\sqrt{n})$$

Combining the last two results, it follows now easily that (20) carries over for $\hat{\theta} \quad \tilde{\theta}/\|\tilde{\theta}\|$. More generally we have:

THEOREM 5. *Let Assumptions 1-3 and hypothesis* $\mathfrak{C}(r)$ *hold, let* $\hat{\theta}$ *and* θ *be the* $k \times r$ *matrices of orthonormal eigenvectors of the matrices* $\hat{M}_1$ *and* M_1*, respectively, corresponding to the r minimum eigenvalues, let* Q_* *be the matrix of the other k-r orthonormal eigenvectors of* M_1*, and let* Λ_* *be the diagonal matrix of corresponding k-r largest eigenvalues. Then*

$$\sqrt{n}(\hat{\theta} \quad \theta) \quad \rightarrow \quad Q_*\Lambda_*^{-1}Q_*^T\left(\int F(x)\overline{W}_r(x)^T dx\right)\left(\theta^T C(1)C(1)^T\theta\right)^{1/2}$$

in distr.

Along the same lines it can be shown that also the other two types of estimators of the co-trending vectors are consistent, but with a lower rate of convergence. Thus the estimator $\hat{\theta}$ of θ on the basis of $\hat{M}_1$ is optimal.

7. COMPARISON WITH COINTEGRATION

The nonlinear trend stationarity hypothesis and the unit root with drift hypothesis are difficult to distinguish. Therefore, we shall also derive the asymptotic distribution of our test under the unit root hypothesis with possible cointegration. Thus, let the data generating process now be:

$$\Delta z_t \quad \beta_1 \quad u_t,$$

where u_t obeys Assumption 2 (but not Assumption 3 of course), hence

$$z_t \quad \beta_0 \quad \beta_1 t \quad \sum_{j=1}^{t} u_j, \quad \textit{where } \beta_0 \quad z_0 .$$

Now denote

$$W_k^*(x) \quad W_k(x) \quad (6x\ \ 4)\int W_k(y)dy \quad (12x\ \ 6)\int yW_k(y)dy,$$

$$W_k^{**}(x) \quad \int_0^x W_k^*(y)dy.$$

Then:

THEOREM 6. *Let* $z_t = z_{t-1} + \beta_1 + u_t$ *be a k-variate unit root with drift process with* u_t *obeying Assumption 2, and let* $\hat{M}_2$ *be defined by* (5) *with* $m = [n^{\alpha}]$ *for some* $\alpha \in (0,1)$. *Suppose there are r cointegrating vectors, and that for each cointegrating vector* θ, $\theta^T D(1)D(1)^T\theta > 0$. *Let* $\hat{\lambda}_1 \le \hat{\lambda}_2 \le .... \le \hat{\lambda}_k$ *be the ordered solutions of the generalized eigenvalue problem* (16). *Then the vector* $n^{1-\alpha}(\hat{\lambda}_1, \dots, \hat{\lambda}_r)$ *converges in distribution to the vector of ordered eigenvalues of the matrix*

$$\int \overline{W}_r(x)\overline{W}_r(x)^T dx \quad \int \overline{W}_r(x)W_{k-r}^{**}(x)^T dx\left(\int W_{k-r}^{**}(x)W_{k-r}^{**}(x)dx\right)^{-1}\int W_{k-r}^{**}(x)\overline{W}_r(x)^T dx,$$

and $(\hat{\lambda}_{r+1}, \dots, \hat{\lambda}_k)$ *converges in distribution to the vector of ordered solutions of the generalized eigenvalue problem*

$$\det\left[\int W_{k-r}^{**}(x)W_{k-r}^{**}(x)^T dx \quad \lambda\int W_{k-r}^{*}(x)W_{k-r}^{*}(x)^T dx\right] \quad 0.$$

Proof: Appendix.

This result shows that our co-trending tests are incapable to distinguish nonlinear co-trending from cointegration. Therefore, before we apply the co-trending tests we should first test the unit root with drift hypothesis against nonlinear trend stationarity. The latter tests are proposed by Bierens (1996a).

8. NONLINEAR CO-TRENDING ANALYSIS OF INTEREST AND INFLATION,

8.1. The data

We use monthly time series of the federal funds rate (FFR) and the CPI inflation rate (CPIR) (i.e., the annual percentage change of the consumer price index), for months 1954.07 through 1994.12. The two series are plotted in Figure 1.

<Insert Figure 1 about here>

We see clearly that these series have a common pattern. The cross-correlation between FFR and lagged CPIR, and CPIR and lagged FFR, is maximal 0.7747 for lag = 0. Thus, the price puzzle is quite apparent. Moreover, it is also clear from Figure 1 that the series do not have a linear trend, or, if they are unit root processes, do not have drift. Therefore our nonlinear co-trending tests can be conducted without detrending.

8.2. Unit root and stationarity test results

First, we have checked whether these series are unit root processes, by conducting the Phillips-Perron (1988) unit root test (PP), Bierens' unit root tests HOAC(1,1) and HOAC(2,2) on the basis of higher order sample autocorrelations, and the Bierens-Guo (1993) tests 1 thought 4, indicated below by BG(1) through BG(4), of the stationarity hypothesis. The first two types of unit root tests test the null hypothesis $\rho \quad 1$ in the auxiliary regression $y_t \quad \rho y_{t-1} \quad \gamma \quad u_t$, where u_t is a zero-mean stationary process for which the functional central limit theorem holds, against the alternative $|\rho| < 1$. The Bierens-Guo (1993) tests reverse the role of this null and alternative. The null distribution of the four tests involved is the absolute value of a standard Cauchy variate. The Phillips-Perron (1988) test and Bierens-Guo (1993) test BG(4) employ a Newey-West (1987) type variance estimator with truncation parameter $m_1 = [cn^r]$, $c > 0$, $0 < r < 1/3$, where n is the sample size. The values used are $c = 5$, $r = .2$. The HOAC(1,1) and HOAC(2,2) tests depend on parameters $c > 0$, $\mu > 0$, and $0 < \delta < 1$, and the lag length is $m_2 = 1 + [cn^{\delta\mu/(3\mu+2)}]$. The values used are $c = 5$, $\mu = 2$, and $\delta = .5$. For both series we have $n = 486$ and thus $m_1 = 17$, $m_2 = 11$. The test results are presented in Table 3.

Table 3: *Unit root and stationarity test results*

Test	*FFR*	*CPIR*	5% *crit. region*	10% *crit. region*	H_0	H_1
PP	11.25	9.98	< 14.00	< 11.20	*U.R*	*Stat.*
HOAC(1,1)	11.59	3.56	< 14.00	< 11.20	*U.R.*	*Stat.*
HOAC(2,2)	11.19	3.77	< 15.70	< 13.10	*U.R.*	*Stat.*
BG(1)	7.78	42.85	>12.71	>6.31	*Stat.*	*U.R.*
BG(2)	8.49	305.37	>12.71	>6.31	*Stat.*	*U.R.*
BG(3)	1.87	2.13	>12.71	>6.31	*Stat.*	*U.R.*
BG(4)	1.39	0.72	>12.71	>6.31	*Stat.*	*U.R.*

The results in Table 3 are mixed. The first three tests reject the unit root hypothesis and the tests BG(3) and BG(4) accept the stationarity hypothesis at the 10% significance level for the FFR, but the test HOAC(2,2) does not reject the unit root hypothesis, and the tests BG(1) and BG(2) accept the stationarity hypothesis only at the 5% level. None of the unit root tests reject the unit root hypothesis for the CPIR, and the tests BG(1) and BG(2) strongly reject the stationarity hypothesis, whereas the tests BG(3) and BG(4) do not reject the stationarity hypothesis. The mixed results for the FFR and the CPIR indicate that these series are neither genuine unit root processes nor genuine stationary processes.

As a double check on whether the FFR and CPIR are unit root processes we have conducted Bierens' (1996b) nonparametric cointegration test to FFR/100 and CPIR/100 separately. This test becomes a unit root test when applied to a single time series. Bierens' (1996b) cointegration test seems to work the best if the variables are in logs. Therefore we have applied the test to the FFR and the CPIR divided by 100, in order to resemble 12 months differences of logs. For both series the test rejected the unit root hypothesis at the 5% significance level.

We also have conducted Bierens' (1996a) tests of the unit root hypothesis against nonlinear trend stationarity, but the results were not conclusive. The latter may be due to the lack of smoothness of the nonlinear trends. Bierens' (1996a) tests are augmented Dickey-Fuller type tests, where the nonlinear trend in the auxiliary regression is represented by a linear trend plus detrended Chebishev time polynomials. The tests allow for a maximum order 20 of the detrended Chebishev polynomials, which for the time series under review may be too low.

Finally, we have applied our nonlinear co-trending test to each of the two time series separately, in order to distinguish between stationarity and nonlinear trend stationarity. Note that our test becomes a test of the null hypothesis of stationarity against the alternative of nonlinear trend stationarity when it is applied to a single time series. The parameter α in Theorem 1 has been chosen equal to 1/2, although a lower value is more favorable as far as the asymptotic power of the test is concerned. As argued before, the value 1/2 is optimal for the convergence of $n^{\alpha}\theta^T\hat{M}_2\theta$ to $\theta^T C(1)C(1)^T\theta$ [cf. (A.10) in the appendix], and that therefore too small an α may cause size distortion. Therefore, the value α 1/2 will be used throughout the empirical application. The values of the test statistics involved are 0.69334 for the FFR and 0.98699 for the CPIR. Comparing these values with the critical values in Table 2, in particular the 5% critical value 0.4657737, we see that for both time series the stationarity hypothesis is rejected at the 5%

significance level in favor of nonlinear trend stationarity.

8.3. Nonlinear co-trending test and estimation results

The components of the vector time series process z_t are now the CPIR and the FFR, for t = 1 (=1954.07) to 486 (=1994.12). The matrices $\hat{M}_1$ and $\hat{M}_2$ are:

$$\hat{M}_1 \quad \begin{pmatrix} 0.2684712 & 0.3321255 \\ 0.3321255 & 0.4580112 \end{pmatrix}, \quad \hat{M}_2 \quad \begin{pmatrix} 8.536301 & 7.509470 \\ 7.509470 & 10.23015 \end{pmatrix},$$

the ordered generalized eigenvalues of $\hat{M}_1$ w.r.t. $\hat{M}_2$ are

$$\hat{\lambda}_1 \quad 0.009134405,\ \hat{\lambda}_2 \quad 0.04478581,$$

and the corresponding standardized generalized eigenvectors of $\hat{M}_1$ w.r.t. $\hat{M}_2$ are

$$\begin{array}{lll} 1 & 0.036827 & \leftarrow CPIR \\ 0.772864 & 1 & \leftarrow FFR \end{array}$$

Multiplying $\hat{\lambda}_r$ by $\sqrt{n} \quad \sqrt{486}$ now yields the test of the null hypothesis that there are r co-trending vectors against the alternative that there are less than r co-trending vectors. The test results presented in Table 4 indicate that there is one co-trending vector.

Table 4: *Tests of the number r of co trending vectors*

r	*test statistic* $\sqrt{n}\hat{\lambda}_r$	10% *crit. region* (*conclusion*)	5% *crit. region* (*conclusion*)
1	0.20137	>0.35182 (*accept*)	>0.46577 (*accept*)
2	0.98732	>0.53561 (*reject*)	>0.67420 (*reject*)

In Figures 2 and 3 we display the components of the estimated functions F and F', respectively, standardized between -1 and 1 by dividing each component by its maximum absolute value. The common patterns in these components clearly corroborate the test result of presence of nonlinear co-trending.

<Insert Figures 2 and 3 about here>

As argued in section 8, the best way of estimating the co-trending vector is to use the eigenvector of the matrix $\hat{M}_1$ corresponding to the smallest eigenvalue. This standardized eigenvector is

$$\begin{array}{ll} 1 & \leftarrow FFR \\ 0.75457 & \leftarrow CPIR \end{array}$$

Thus:

Nonlinear trend in CPIR = 0.75457 x *Nonlinear trend in FFR*

Note that the significance of the parameter involved is already established, by testing whether each of these time series are stationary. Moreover, note that the way we have written the nonlinear co-trending relations should not be interpreted as a causal ordering, as each of the nonlinear trends in FFR and CPIR may be considered to be the common nonlinear trend.

In order to determine the estimation error of the estimate 0.75457, we have tested the hypothesis that the vector $H = (1, -a)^T$ is a co-trending vector. The results are presented in Table 5 for a ranging from 0.3 to 1.2.

Table 5: *Test of the hypothesis that*
H (1, $a)^T$ *is a co trending vector*

	test	conclusion	
a	*statistic*	10%	5%
1.2	0.55022	*reject*	*reject*
1.1	0.46153	*reject*	*accept*
1	0.36609	*reject*	*accept*
0.9	0.27766	*accept*	*accept*
0.8	0.21696	*accept*	*accept*
0.7	0.20276	*accept*	*accept*
0.6	0.23919	*accept*	*accept*
0.5	0.31274	*accept*	*accept*
0.4	0.40249	*reject*	*accept*
0.3	0.49162	*reject*	*reject*

Thus the 95% confidence interval of the the parameter a is approximately (0.3, 1.2), and the 90% confidence interval is approximately (0.4, 1).

Summarizing, our findings suggest that the empirical phenomenon known as the *price puzzle* (i.e. the positive correlation between the inflation rate and the interest rate), is due to a common nonlinear deterministic time trend.

APPENDIX

Proof of Lemma 1: It follows from (6), (7) and the first part of (8) that

$$\sum_{j=1}^{n-1} j^4 \|\gamma_{j,n}\|^2 \quad O(1).$$

which, by Liapounov' inequality, implies that

$$\text{(A.1)} \qquad \sum_{j=1}^{n-1} j\|\gamma_{j,n}\| \leq \left(\sum_{j=1}^{n-1} j^4 \|\gamma_{j,n}\|^2\right)^{1/2} \left(\sum_{j=1}^{n-1} j^{-2}\right)^{1/2} \quad O(1),$$

and thus

$$\text{(A.2)} \qquad \sum_{j=1}^{n-1} \|\gamma_{j,n}\| \quad O(1).$$

Similarly, it follows from Assumption 1 and the orthonormality of the functions $\surd 2\cos(j\pi x)$ w.r.t. the Lebesgue measure on [0,1] that

$$\sum_{j=1}^{\infty} j^4 \|\gamma_j\|^2 < \infty, \quad \sum_{j=1}^{\infty} j\|\gamma_j\| < \infty.$$

and

$$\text{(A.3)} \qquad \sum_{j=1}^{n-1} \|\gamma_{j,n} \quad \gamma_j\|^2 \quad O(1).$$

Next, observe that

$$F_n(x) \quad (1/n)\sum_{j=1}^{n-1} \gamma_{j,n}\sqrt{2}\sum_{t=1}^{[nx]} \cos(j\pi(t \;\; .5)/n)$$

$$(1/n)\sum_{j=1}^{n-1} \gamma_{j,n}\frac{1}{\sqrt{2}}\sum_{t=1}^{[nx]} \left(\exp(ij\pi/n)\right)^t \exp(.5ij\pi/n))$$

$$(1/n)\sum_{j=1}^{n-1} \gamma_{j,n}\frac{1}{\sqrt{2}}\sum_{t=1}^{[nx]} \left(\exp(\;\; ij\pi/n)\right)^t \exp(\;\; .5ij\pi/n))$$

$$\frac{\sqrt{2}}{\pi}\sum_{j=1}^{n-1} \frac{\gamma_{j,n}}{j}\sin(j\pi[nx]/n)\left(\frac{\sin(.5j\pi/n)}{.5j\pi/n}\right)^{-1}.$$

Since for j = 1,...,n-1 and n = 1,2,.....,

$$\frac{2}{\pi} < \frac{\sin(.5j\pi/n)}{.5j\pi/n} < 1,$$

it follows from (A.1) and (A.2) that

$$F_n(x) \quad \frac{\sqrt{2}}{\pi}\sum_{j=1}^{n-1}\frac{\gamma_{j,n}}{j}\sin(j\pi x)\left(\frac{\sin(.5j\pi/n)}{.5j\pi/n}\right)^{-1} \quad O(1/n)$$

(A.4)
$$\frac{\sqrt{2}}{\pi}\sum_{j=1}^{n-1}\frac{\gamma_{j,n}}{j}\sin(j\pi x) \quad O\left(\sum_{j=1}^{n-1}\frac{|\gamma_{j,n}|}{j}\left(1 \quad \frac{\sin(.5j\pi/n)}{.5j\pi/n}\right)\right) \quad O(1/$$

$$\frac{\sqrt{2}}{\pi}\sum_{j=1}^{n-1}\frac{\gamma_{j,n}}{j}\sin(j\pi x) \quad O(1/n)$$

Lemma 1 follows now easily from (A.3), (A.4), (10), and the orthonormality of the functions $\sqrt{2}\sin(j\pi x)$ w.r.t. the Lebesgue measure on [0,1].

Proof of Lemma 2: Since (A.4) holds uniformly on [0,1], we have

$$\frac{(x \quad \delta_n) \quad F_n(x)}{\delta_n} \quad \frac{\sqrt{2}}{\pi}\sum_{j=1}^{n-1}\frac{\gamma_{j,n}}{j}\frac{\sin(j\pi(x \ \delta_n) \quad \sin(j\pi x)}{\delta_n} \quad O\left(\frac{1}{\delta_n}\right.$$

(A.5)
$$\sum_{j=1}^{n-1}\gamma_{j,n}\sqrt{2}\cos(j\pi x) \quad O\left(\delta_n\sum_{j=1}^{n-1}j\|\gamma_{j,n}\|\right) \quad O\left(\frac{1}{\delta_n n}\right)$$

$$\phi_n(x) \quad O(\delta_n) \quad O\left(\frac{1}{\delta_n n}\right)$$

Proof of Lemma 3, Part (14): This part follows easily from Lemma 1 and (13).

Proof of Lemma 3, Part (15): We prove (15) only for the scalar case $z_t \in \mathbb{R}$. Let m be a natural number between zero and n, depending on the sample size n such that $m \to \infty$ and $m/n \to 0$. It follows from (11) and the easy results

$$\sqrt{n}(\hat{\beta}_0 \ \beta_{0,n}) \quad O_p(1), \ n\sqrt{n}(\hat{\beta}_1 \ \beta_{1,n}) \quad O_p(1)$$

that for $t \geq m$,

$$\overset{def.}{\quad} (1/m)\sum_{j=0}^{m-1}\left(z_{t-j} \quad \hat{\beta}_0 \quad \hat{\beta}_1(t\ j)\right) \quad C(1)(1/m)\sum_{j=0}^{m-1}\varepsilon_{t-j} \quad (v_t \quad v_{t-r}$$

$$(\hat{\beta}_0\ \beta_{0,n}) \quad (\hat{\beta}_1\ \beta_{1,n})\left(t\ \tfrac{1}{2}(m\ 1)\right) \quad (1/m)\sum_{j=0}^{m-1} f_n(t\ j) \tag{A.6}$$

$$C(1)(1/m)\sum_{j=0}^{m-1}\varepsilon_{t-j} \quad (1/m)\sum_{j=0}^{m-1} f_n(t\ j) \quad (v_t \quad v_{t-m})/m \quad O_p(1/\sqrt{n}$$

where the last term in (A.6) is uniform in $t = m,..,n$. Moreover, it follows from (2) and Lemma 2 that

$$(1/m)\sum_{j=0}^{m-1} f_n(t\ j) \quad \frac{F_n\left(\frac{t}{n}\right) \quad F_n\left(\frac{t}{n}\ \frac{m}{n}\right)}{\frac{m}{n}} \tag{A.7}$$

$$\phi_n\left(\frac{t\ m}{n}\right) \quad O(m/n) \quad O(1/m),$$

where the last two terms are uniform in $t = m,..,n$. Furthermore, observe from (7) and (A.1) that $\|\phi_n(x)\| < \sqrt{2}\Sigma_{j=1}^{n-1}\|\gamma_{j,n}\| < \infty$, hence

$$(1/n)\sum_{t=m}^{n}\phi_n((t\ m)/n)^2 \quad (1/n)\sum_{t=1}^{n}\phi_n(t/n)^2 \quad O(m/n). \tag{A.8}$$

Then it follows from Assumption 1, (5), (A.6) through (A.8), and Lemmas A.1-2 below that

$$\hat{M}_2 \quad (1/n)\sum_{t=m}^{n}\eta_t\eta_t^T \quad \int F'(x)F'(x)^T dx \quad o_p(1).$$

LEMMA A.1: *Under Assumption 1,* $(1/n)\Sigma_{t=1}^{n}\phi_n(t/n)\phi_n(t/n)^T \quad \int\phi_n(x)\phi_n(x)^T dx \quad O(1/n)$.

LEMMA A.2: *Let* ε_t *be a sequence of independent standard normally distributed random variabled. Then* $(1/n)\Sigma_{t=m}^{n}\left((1/\sqrt{m})\Sigma_{j=0}^{m-1}\varepsilon_{t-j}\right)^2 \quad 1 \quad O_p(m/n)$.

Proof of Lemma A.1: Without loss of generality we may assume $z_t \in \mathbb{R}$. Then it follows from (6) and (7) that

$$(1/n)\sum_{t=1}^{n}\phi_n(t/n)^2 \quad (1/n)\sum_{t=1}^{n}\left(\sqrt{2}\sum_{j=1}^{n-1}\gamma_{j,n}\cos(j\pi(t\;\;0.5)/n)\right)^2 \quad \sum_{j=1}^{n-1}\gamma_{j,n}^2$$

$$\int\left(\sqrt{2}\sum_{j=1}^{n-1}\gamma_{j,n}\cos(j\pi x)\right)^2 dx \quad \int\left(\phi_n(x) \quad O(1/n)\right)^2 dx \quad \int\phi_n(x)^2 dx \quad O(1/n).$$

Proof of Lemma A.2: Let $N = [n/m]$. Then

$$(1/n)\sum_{t=m}^{n}\left((1/\sqrt{m})\sum_{j=1}^{m-1}\varepsilon_{t-j}\right)^2 \quad \frac{m(N\;\;1)}{n}\frac{1}{m}\sum_{i=0}^{m-1}\left[\frac{1}{N\;\;1}\sum_{\tau=1}^{N-1}\left((1/\sqrt{m})\sum_{j=1}^{m-1}\varepsilon_{\tau m+i-j}\right)^2\right]$$

$$O_p\left(\frac{n\;\;mN}{n}\right) \quad \frac{m(N\;\;1)}{n}\frac{1}{m}\sum_{i=0}^{m-1}\left[1 \quad O_p(1/(N\;\;1))\right] \quad O_p\left(\frac{n\;\;mN}{n}\right) \quad 1 \quad O_p(m/n)$$

due to the fact the expression between square brackets in the first line is a mean of independent $\chi^2(1)$ distributed random variables.

Proof of Lemma 4: It follows from (13) that

$$\begin{pmatrix}\sqrt{n} & 0^T\\ 0 & I_{k-1}\end{pmatrix} Q^{\,T}\hat{M}_1 Q \begin{pmatrix}\sqrt{n} & 0^T\\ 0 & I_{k-1}\end{pmatrix} \quad \begin{pmatrix} n\theta^T\hat{M}_1\theta & \sqrt{n}\theta^T\hat{M}_1 Q_* \\ \sqrt{n}Q_*^{\,T}\hat{M}_1\theta & Q_*^{\,T}\hat{M}_1 Q_*\end{pmatrix}$$

(A.9)

$$\begin{pmatrix}\theta^T C(1)\int\overline{W}_k(x)\overline{W}_k(x)^T dx C(1)^T\theta & \theta^T C(1)^T\int\overline{W}_k(x)F(x)^T dx Q_* \\ Q_*^{\,T}\int F(x)\overline{W}_k(x)^T dx C(1)^T\theta & \Lambda_*\end{pmatrix}$$

in distr. Now observe from part (9) of Assumption 1 that under the hypothesis $\mathfrak{C}$ (1), rank of the matrix M_1 is k-1, hence Λ_* is nonsingular. Inverting the partitioned matrices in (A.9), the lemma follows.

Proof of Lemma 5: Since $\theta^T f_n(t) \quad 0$, it follows from (A.6) and Lemma A.2 that

$$
\begin{aligned}
{}^{T}\hat{M}_2\theta \quad & \frac{1}{n}\sum_{t=m}^{n}\left(\theta^T C(1)\frac{1}{\sqrt{m}}\sum_{j=0}^{m-1}\varepsilon_{t-j} \quad \theta^T(v_t \quad v_{t-m})/\sqrt{m} \quad O_p(\sqrt{m/}\right. \\
& \times\left(\theta^T C(1)\frac{1}{\sqrt{m}}\sum_{j=0}^{m-1}\varepsilon_{t-j} \quad \theta^T(v_t \quad v_{t-m})/\sqrt{m} \quad O_p(\sqrt{}\right. \\
& \theta^T C(1)C(1)^T\theta \quad O_p(1/\sqrt{m}) \quad O_p(\sqrt{m/n}),
\end{aligned}
\tag{A.10}
$$

where the rate of convergence is optimal for m proportional to $[\sqrt{n}]$. Furthermore, similarly to (A.10) it can be shown that

$$\sqrt{m}\hat{M}_2\theta \quad o_p(1). \tag{A.11}$$

Denoting again by $Q = (\theta, Q_*)$ the orthogonal matrix of eigenvectors of the matrix M_1, it follows now similarly to (A.9) that

$$\begin{pmatrix} \sqrt{m} & 0^T \\ 0 & I_{k-1} \end{pmatrix} Q^T\hat{M}_2 Q \begin{pmatrix} \sqrt{m} & 0^T \\ 0 & I_{k-1} \end{pmatrix} \rightarrow \begin{pmatrix} \theta^T C(1)C(1)^T\theta & 0^T \\ 0 & Q_*^T M_2 Q_* \end{pmatrix} \tag{A.12}$$

in prob. Therefore, the minimum eigenvalue of the matrix $\hat{M}_2$, times m, converges in probability to $\theta^T C(1)C(1)^T\theta$. Since by part (9) of Assumption 1, $\operatorname{rank}\left[\int F'(x)F'(x)^T dx\right] \quad k \;\, 1$ under $\mathfrak{C}(1)$, and consequently $Q_*^T M_2 Q$ is nonsingular, it follows now from the result of Anderson, Brons and Jensen (1983), and from (A.12), that Lemma 5 holds.

Proof of Lemma 6: The proof employs Mercer's theorem. Cf. Dunford and Schwartz (1963, p.1088), and Bierens and Ploberger (1995). Let

$$\Gamma(x,y) \quad E\left(\bar{W}_1(x)\bar{W}_1(y)\right).$$

This function is real valued symmetric positive semi-definite, and it follows from (13) that Γ is continuous on $[0,1]\times[0,1]$. Now Mercer's theorem states that there exists a sequence λ_j of nonnegative eigenvalues and corresponding sequence $\psi_j(x)$ of real valued continuous eigenfunctions such that:

$$\int \Gamma(x,y)\psi_j(y)dy \quad \lambda_j\psi_j(x),\ j \quad 1,2,......,$$

$$\sum_{j=1}^{\infty}\lambda_j < \infty,\quad \int \psi_i(x)\psi_j(x)dx \quad I(i\ \ j),$$

$$\Gamma(x,y) \quad \sum_{j=1}^{\infty}\lambda_j\psi_j(x)\psi_j(y).$$

Since

$$\int \overline{W}_1(x)dx \quad 0\ \ a.s.,\quad \int x\overline{W}_1(x)dx \quad 0\ \ a.s.,$$

we have that

$$\int \Gamma(x,y)dy \quad 0,\quad \int \Gamma(x,y)ydy \quad 0,$$

hence 1 and (x-½)√12 are eigenfunctions corresponding to zero eigenvalues. Therefore, for all other eigenfunctions we have by orthogonality, $\int \psi_j(x)dx = 0$ and $\int x\psi_j(x)dx = 0$. Moreover, the eigenfunction $\psi_j(x)$, except 1 and (x-½)√12, can be chosen such that $\psi_j(0) = \psi_j(1) = 0$, by using the transformation

$$\psi_j^*(x) \quad \frac{\psi_j(x) \quad \psi_j(0) \quad \left(\psi_j(1)\ \ \psi_j(0)\right)x}{\left[\int\left(\psi_j(x) \quad \psi_j(0) \quad \left(\psi_j(1)\ \ \psi_j(0)\right)x\right)^2dx\right]^{1/2}}.$$

We can now find sub-sequences $j_m(1),...,j_m(k\text{-}1)$ such that for $i,i_1,i_2 = 1,..,k\text{-}1$,

$$\int \psi_{j_m(i_1)}(x)\psi_{j_m(i_2)}(x)dx \quad I(i_1\ \ i_2),$$

$$\int \psi_{j_m(i)}(x)dx \quad 0,\ \int x\psi_{j_m(i)}(x)dx \quad 0,\ \psi_{j_m(i)}(0) \quad \psi_{j_m(1)}(1) \quad 0,$$

$$\lim_{m\to\infty}\max_{1\le i\le k-1}\lambda_{j_m(i)} \quad 0.$$

hence, denoting

$$F_m(x) \quad Q_*\Lambda_*^{1/2}\begin{pmatrix}\psi_{j_m(1)}(x)\\ : \\ \psi_{j_m(k-1)}(x)\end{pmatrix}$$

we have $\int F_m(x)F_m(x)^T dx = M_1$, $F_m(0) = F_m(1) = 0$, $\int F_m(x)dx = 0$, and

$$\lim_{m\to\infty} E\left[\int \overline{W}_1(x)F_m(x)^T dx Q_* \Lambda_*^{-1} Q_*^T \int \overline{W}_1(y)F_m(y)dy\right]$$

$$\lim_{m\to\infty}\sum_{i=1}^{k-1} \int\int \psi_{j_m(i)}(x)\Gamma(x,y)\psi_{j_m(i)}(y)dxdy \quad \lim_{m\to\infty}\sum_{i=1}^{k-1} \lambda_{j_m(i)} \quad 0.$$

The lemma follows now from Chebishev's inequality.

Proof of Theorem 6: Observe that

$$\frac{1}{\sqrt{n}}(\hat{\beta}_0 \quad \beta_0) \quad 4(1/n)\sum_{t=1}^{n} U_n(t/n) \quad 6(1/n)\sum_{t=1}^{n} (t/n)U_n(t/n) \quad o_p(1),$$

$$\sqrt{n}(\hat{\beta}_1 \quad \beta_1) \quad 6(1/n)\sum_{t=1}^{n} U_n(t/n) \quad 12(1/n)\sum_{t=1}^{n} (t/n)U_n(t/n) \quad o_p(1),$$

hence

$$\frac{z_{[nx]} \quad \hat{\beta}_0 \quad \hat{\beta}_1[nx]}{\sqrt{n}} \quad U_n(x) \quad \frac{1}{\sqrt{n}}(\hat{\beta}_0 \; \beta_0) \quad \sqrt{n}(\hat{\beta}_1 \; \beta_1)\frac{[nx]}{n}$$

$$U_n(x) \quad 4(1/n)\sum_{t=1}^{n} U_n(t/n) \quad 6(1/n)\sum_{t=1}^{n} (t/n)U_n(t/n)$$

(A.13)

$$6x(1/n)\sum_{t=1}^{n} U_n(t/n) \quad 12x(1/n)\sum_{t=1}^{n} (t/n)U_n(t/n) \quad o_p(1),$$

$$\Rightarrow C(1)\left(W_k(x) \quad (6x \;\; 4)\int W_k(y)dy \quad (12x \;\; 6)\int y W_k(y)dy\right)$$

$$C(1)W_k^*(x).$$

Thus,

$$\hat{F}(x)/\sqrt{n} \Rightarrow C(1)\int_0^x W_k^*(y)dy \quad C(1)W_k^{**}(x),$$

and consequently

(A.14) $$\frac{1}{n}\hat{M}_1 \to C(1)\int W_k^{**}(x)W_k^{**}(x)^T dx C(1)^T$$

in distr.

Now suppose that z_t is cointegrated with one cointegrating vector θ. Then $\theta^T C(1) = 0^T$. Since by (11), $z_t \quad \beta_0 \quad \beta_1 t \quad C(1)\Sigma_{j=1}^{t}\varepsilon_j \quad v_t \quad v_0$, with $v_t \quad D(L)\varepsilon_t$, we how have

$$\theta^T z_t \quad \theta^T(\beta_0 \; v_0) \quad \theta^T\beta t \quad \theta^T v_t \quad \theta^T(\beta_0 \; v_0) \; \theta^T\beta t \quad \theta^T D(L)\varepsilon_t,$$

and consequently

$$\sqrt{n}\theta^T\hat{F}(x) \Rightarrow \theta^T D(1)\bar{W}_k(x).$$

Let $Q = (\theta, Q_*)$ be the orthogonal matrix of eigenvectors of $C(1)C(1)^T$, corresponding to the increasingly ordered eigenvalues. It is now easy to verify that under the unit root hypothesis with one cointegrated vector,

$$\frac{1}{n}\begin{pmatrix} n & 0^T \\ 0 & I_{k-1} \end{pmatrix} Q^T \hat{M}_1 Q \begin{pmatrix} n & 0^T \\ 0 & I_{k-1} \end{pmatrix} \rightarrow$$

$$\begin{pmatrix} \theta^T D(1)\int \bar{W}_k(x)\bar{W}_k(x)^T dx D(1)^T\theta & \theta^T D(1)\int \bar{W}_k(x) W_k^{**}(x)^T dx C(1)^T Q_* \\ Q_*^T C(1)\int W_k^{**}(x)\bar{W}_k(x)^T D(1)^T\theta & Q_*^T C(1)\int W_k^{**}(x) W_k^{**}(x)^T dx C(1)^T Q_* \end{pmatrix}$$

in distr. Therefore, similarly to Lemma 4, it follows that for every nonnegative sequence $m = o(n)$ we have:

(A.15)
$$\begin{pmatrix} 1/\sqrt{n} & 0^T \\ 0 & \sqrt{n}\sqrt{m/n} I_{k-1} \end{pmatrix} Q^T \hat{M}_1^{-1} Q \begin{pmatrix} 1/\sqrt{n} & 0^T \\ 0 & \sqrt{n}\sqrt{m/n} I_{k-1} \end{pmatrix}$$

$$\begin{pmatrix} 1/\sqrt{m} & 0^T \\ 0 & \sqrt{n} I_{k-1} \end{pmatrix} Q^T \frac{m}{n}\hat{M}_1^{-1} Q \begin{pmatrix} 1/\sqrt{m} & 0^T \\ 0 & \sqrt{n} I_{k-1} \end{pmatrix} \rightarrow \begin{pmatrix} \tilde{\mu}_*^{-1} & 0^T \\ 0 & O \end{pmatrix}$$

in distr., where

$$\tilde{\mu}_{*} \quad \theta^T D(1)\int \overline{W}_k(x)\overline{W}_k(x)^T dx D(1)^T\theta$$

$$\theta^T D(1)\int \overline{W}_k(x)W_k^{**}(x)^T dx C(1)^T Q_{*}\left(Q_{*}^{\,T}C(1)\int W_k^{**}(x)W_k^{**}(x)dx C(1)^T Q_{*}\right)^{-1}$$

$$\times \; Q_{*}^{\,T}C(1)\int W_k^{**}(x)\overline{W}_k(x)^T dx D(1)^T\theta$$

$$\sim \; \theta^T D(1)D(1)^T\theta$$

$$\times \left(\int\left(\overline{W}_1(x)\right)^2 dx \quad \int \overline{W}_1(x)W_{k-1}^{**}(x)^T dx\left(\int W_{k-1}^{**}(x)W_{k-1}^{**}(x)dx\right)^{-1}\int W_{k-1}^{**}(x)\overline{W}_1(x)dx\right)$$

Next we investigate the asymptotic properties of the matrix $\hat{M}_2$ under the unit root with cointegration hypothesis. Let m be as in (A.6). Then it follows from (A.13) that for $x \in [0,1]$,

$$\frac{1}{m\sqrt{n}}\sum_{j=0}^{m-1}\left(z_{[nx]-j} \quad \hat{\beta}_0 \quad \hat{\beta}_1([nx]\;\; j)\right)$$

$$\frac{1}{m}\sum_{j=0}^{m-1} U_n\left(\frac{[nx]\;\; j}{n}\right) \quad 4(1/n)\sum_{t=1}^{n} U_n(t/n) \quad 6(1/n)\sum_{t=1}^{n}(t/n)U_n(t/n)$$

$$6\frac{[nx]}{n}(1/n)\sum_{t=1}^{n} U_n(t/n) \quad 12\frac{[nx]}{n}(1/n)\sum_{t=1}^{n}(t/n)U_n(t/n) \quad o_p(1).$$

Moreover,

$$\frac{1}{m}\sum_{j=0}^{m-1} U_n\left(\frac{[nx]\;\; j}{n}\right) \quad \frac{1}{m}\sum_{j=0}^{m-1} U_n\left(\frac{[nx]\;\; m\;\; 1\;\; j}{n}\right) \quad \frac{1}{m}\sum_{j=0}^{m-1}\int_j^{j+1} U_n\left(\frac{[nx]\;\; m\;\; 1\;\; y}{n}\right)dy$$

$$\frac{1}{m}\int_0^m U_n\left(\frac{[nx]\;\; m\;\; 1}{n} \quad \frac{m}{n}\frac{y}{m}\right)dy \quad \int_0^1 U_n\left(\frac{[nx]}{n} \quad \frac{m\;\; 1}{n} \quad \frac{m}{n}y\right)dy \quad U_n(x) \quad o_p(1)$$

Therefore

$$\frac{1}{m\sqrt{n}}\sum_{j=0}^{m-1}\left(z_{[nx]-j} \quad \hat{\beta}_0 \quad \hat{\beta}_1([nx]\;\; j)\right) \Rightarrow C(1)W_k^{*}(x).$$

and consequently,

$$\text{(A.16)} \qquad \frac{\hat{M}_2}{n} \rightarrow C(1)\int W_k^{*}(x)W_k^{*}(x)^T dx C(1)^T$$

in distr. Furthermore, it is easy to verify that (A.10) goes through, with $C(1)$ replaced by $D(1)$:

$$\frac{m}{n}\theta^T\hat{M}_2\theta \quad \theta^T D(1)D(1)^T\theta \quad o_p(1),$$

and that (A.11) now reads:

$$\frac{\sqrt{m}}{n}\hat{M}_2\theta \quad o_p(1).$$

Thus,

$$\begin{pmatrix} \sqrt{m} & 0^T \\ 0 & \frac{1}{\sqrt{n}}I_{k-1} \end{pmatrix} Q^T\hat{M}_2 Q \begin{pmatrix} \sqrt{m} & 0^T \\ 0 & \frac{1}{\sqrt{n}}I_{k-1} \end{pmatrix} \to$$

$$\begin{pmatrix} \theta^T D(1)D(1)^T\theta & 0^T \\ 0 & Q_*^T C(1)\int W_k^*(x)W_k^*(x)^T dx C(1)^T Q_* \end{pmatrix}$$

in distr, and therefore, similarly to Lemma 5, we have

$$\begin{pmatrix} 1/\sqrt{m} & 0^T \\ 0 & \sqrt{n}I_{k-1} \end{pmatrix} Q^T\hat{M}_2^{-1} Q \begin{pmatrix} 1/\sqrt{m} & 0^T \\ 0 & \sqrt{n}I_{k-1} \end{pmatrix}$$

(A.17)

$$\to \begin{pmatrix} \left(\theta^T D(1)D(1)^T\theta\right)^{-1} & 0^T \\ 0 & \left(Q_*^T C(1)\int W_k^*(x)W_k^*(x)^T dx C(1)^T Q_*\right)^{-1} \end{pmatrix}$$

in distr.

Comparing (A.15) and (A.17) with eigenvalue problem (17), we see that under the unit root hypothesis with single cointegration and $m = [n^\alpha]$, with $0 < \alpha < 1$, the minimum solution $\hat{\lambda}_1$ of eigenvalue problem (16) satisfies:

$$n^{1-\alpha}\hat{\lambda}_1 \to \frac{\tilde{\mu}_*}{\theta^T D(1)D(1)^T\theta}$$

$$\int\left(\bar{W}_1(x)\right)^2 dx \quad \int \bar{W}_1(x)W_{k-1}^{**}(x)^T dx\left(\int W_{k-1}^{**}(x)W_{k-1}^{**}(x)dx\right)^{-1}\int W_{k-1}^{**}(x)\bar{W}_1(x)dx$$

in distr. Note that the limiting random variable involved has the same upperbound as in the case of co-trending. Moreover, it follows from (A.14) and (A.16) that under the unit root hypothesis without cointegration, $\hat{\lambda}_1$ converges in distribution to the minimum solution of the generalized eigenvalue problem

$$\det\left[\int W_{k-1}^{**}(x)W_{k-1}^{**}(x)^T dx \quad \lambda\int W_{k-1}^{*}(x)W_{k-1}^{*}(x)^T dx\right] \quad 0.$$

This completes the proof of Theorem 6 for the case of one cointegrating vector. The general case can be shown along similar lines.

REFERENCES

ANDERSON, S.A., H.K.BRONS and S.T.JENSEN (1983): "Distribution of Eigenvalues in Multivariate Statistical Analysis", *Annals of Statistics* 11, 392-415.

BALKE, N.S. and K.M.EMERY (1994): "The Federal Funds Rate as an Indicator of Monetary Policy: Evidence from the 1980's", *Federal Reserve Bank of Dallas Economic Review*, 1-15

BALKE, N.S. and K.M.EMERY (1995): "Understanding the Price Puzzle", Working paper, Department of Economics, Southern Methodist University

BERNANKE, B.S. and A.S.BLINDER (1992): "The Federal Funds Rate and the Channels of Monetary Transmission", *American Economic Review* 82, 901-921.

BIERENS, H.J. (1993): "Higher-order Sample Autocorrelations and the Unit Root Hypothesis", *Journal of Econometrics* 57, 137-160.

BIERENS, H.J. (1994): *Topics in Advanced Econometrics: Estimation, Testing, and Specification of Cross-Section and Time Series Models* (Cambridge, U.K.: Cambridge University Press).

BIERENS, H.J. (1996a): "Testing the Unit Root with Drift Hypothesis Against Nonlinear Trend Stationarity, with an Application to the U.S. Price Level and Interest Rate", *Journal of Econometrics* (forthcoming).

BIERENS, H.J. (1996b): "Nonparametric Cointegration Analysis", *Journal of Econometrics* (forthcoming).

BIERENS, H.J. and S.GUO (1993): "Testing Stationarity and Trend Stationarity Against the Unit Root Hypothesis", *Econometric Reviews* 12, 1-32.

BIERENS, H.J. and W.PLOBERGER (1995): "Asymptotic Theory of Integrated

Conditional Moment Tests", working paper, Department of Economics, Southern Methodist University, Dallas, and CentER, Tilburg University, the Netherlands.

BILLINGSLEY, P. (1968): *Convergence of Probability Measures* (New York: John Wiley).

CHRISTIANO, L.J. and M.EICHENBAUM (1992), "Identification and the Liquidity Effect of Monetary Policy Shocks", in A.Cuikerman, Z.Hercowitz and L.Leiderman (Eds), *Political Economy, Growth, and Business Cycles.* (Cambridge, Massachusetts: M.I.T. Press)

CHRISTIANO, L.J., M.EIGENBAUM, and C.EVANS (1994): "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds", Federal Reserve Bank of Chicago Working Paper Series 94-2

CHRISTIANO, L.J., M.EICHENBAUM, and C.EVANS (1995): "Identification and the Effect of Monetary Policy Shocks", in: M.Blejer, Z.Eckstein, Z.Hercowitz, and L.Leiderman (Eds), *Factors in Economic Stabilization and Growth.* (Cambridge, U.K.: Cambridge University Press)

DUNFORD, N. and J.T.SCHWARTZ (1963): *Linear Operators, Part II: Spectral Theory* (New York: Wiley Interscience).

EICHENBAUM, M. (1992)," Comment on 'Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy' by Christopher Sims", *European Economic Review* 36, 1001-1011

ENGLE, R.F. (1987): "On the Theory of Cointegrated Economic Time Series", Invited paper presented at the Econometric Society European Meeting 1987, Copenhagen.

HALL, P. and C.C.HEYDE (1980): *Martingale Limit Theory and Its Applications* (San Diego: Academic Press).

HAMMING, R.W. (1973): *Numerical Methods for Scientists and Engineers* (New York: Dover Publications).

JOHANSEN, S. (1988): "Statistical Analysis of Cointegrated Vectors", *Journal of Economic Dynamics and Control* 12, 231-254.

JOHANSEN, S. (1991): "Estimation and Hypothesis Testing of Cointegrated Vectors in Gaussian Vector Autoregressive Models", *Econometrica* 59, 1551-1580.

JOHANSEN, S. (1994): "The Role of the Constant and Linear Terms in Cointegration Analysis of Nonstationary Variables", *Econometric Reviews* 13, 205-230.

JOHANSEN, S. and K.JUSELIUS (1990): "Maximum Likelihood Estimation and Inference on Cointegration, with Applications to the Demand for Money", *Oxford Bulletin of Economics and Statistics* 52, 169-210.

NELSON, C.R. and C.I. PLOSSER (1982): "Trends and Random Walks in Macro-economic Time Series", *Journal of Monetary Economics* 10, 139-162.

NEWEY, W.K. and K.D.WEST (1987): "A Simple Positive Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix", *Econometrica* 55, 703-708

PERRON, P. (1988): "Trends and Random Walks in Macroeconomic Time Series: Further Evidence from a New Approach", *Journal of Economic Dynamics and Control* 12, 297-332.

PERRON, P. (1989): "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis", *Econometrica* 57, 1361-1402.

PERRON, P. (1990): "Testing the Unit Root in a Time Series with a Changing Mean", *Journal of Business and Economic Statistics* 8, 153-162.

PHILLIPS, P.C.B. and P.PERRON (1988), "Testing for a Unit Roots in Time Series Regression", *Biometrika* 75, 335-346.

PHILLIPS, P.C.B. and V.SOLO (1992): "Asymptotics for Linear Processes", *Annals of Statistics* 20, 971-1001.

SCHOTMAN, P.C. and H.K. VAN DIJK (1991): "On Bayesian Routes to Unit Roots", *Journal of Applied Econometrics 6*, 387-401.

SIMS, C.A. (1992), "Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy", *European Economic Review* 36, 975-1000.

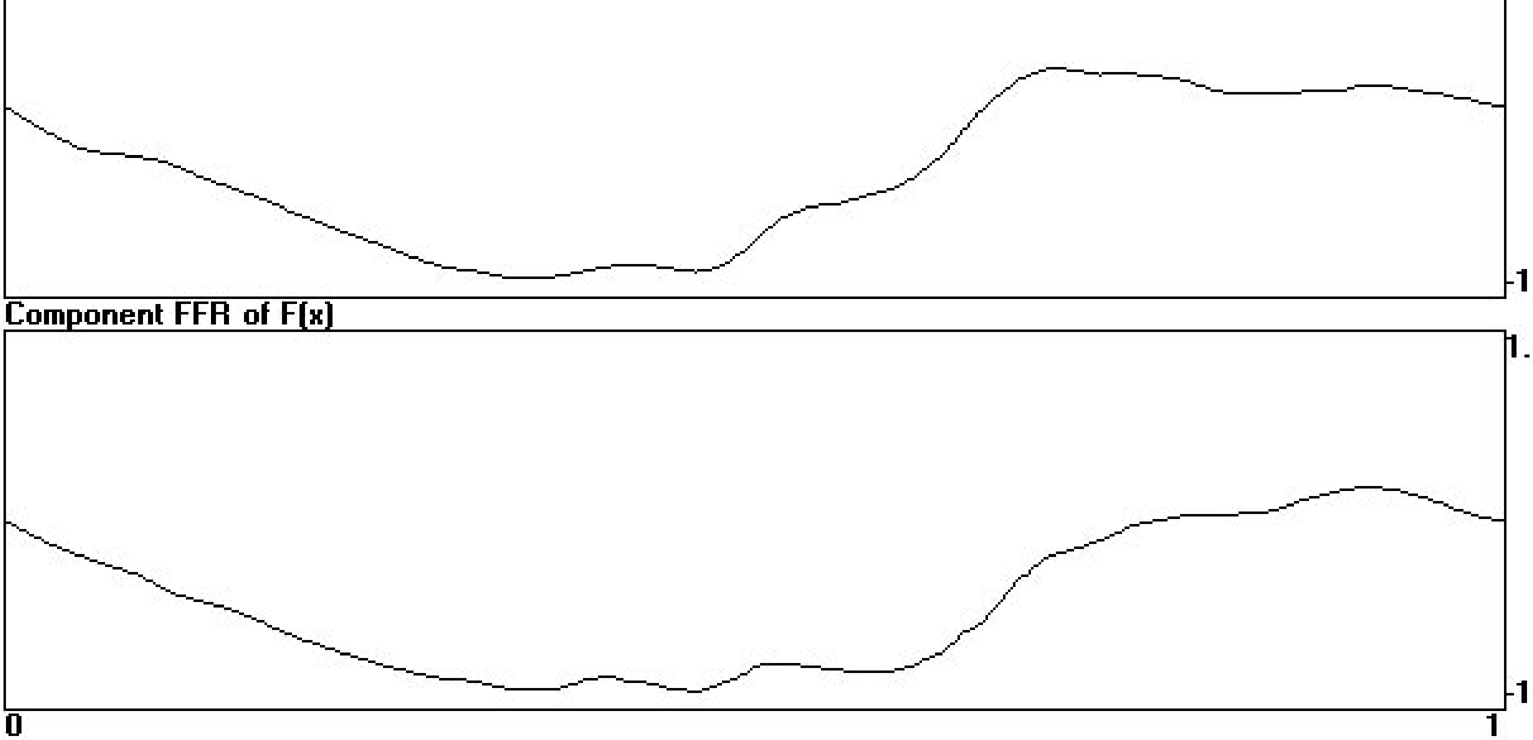

Figure 2: Rescaled components of F(x)

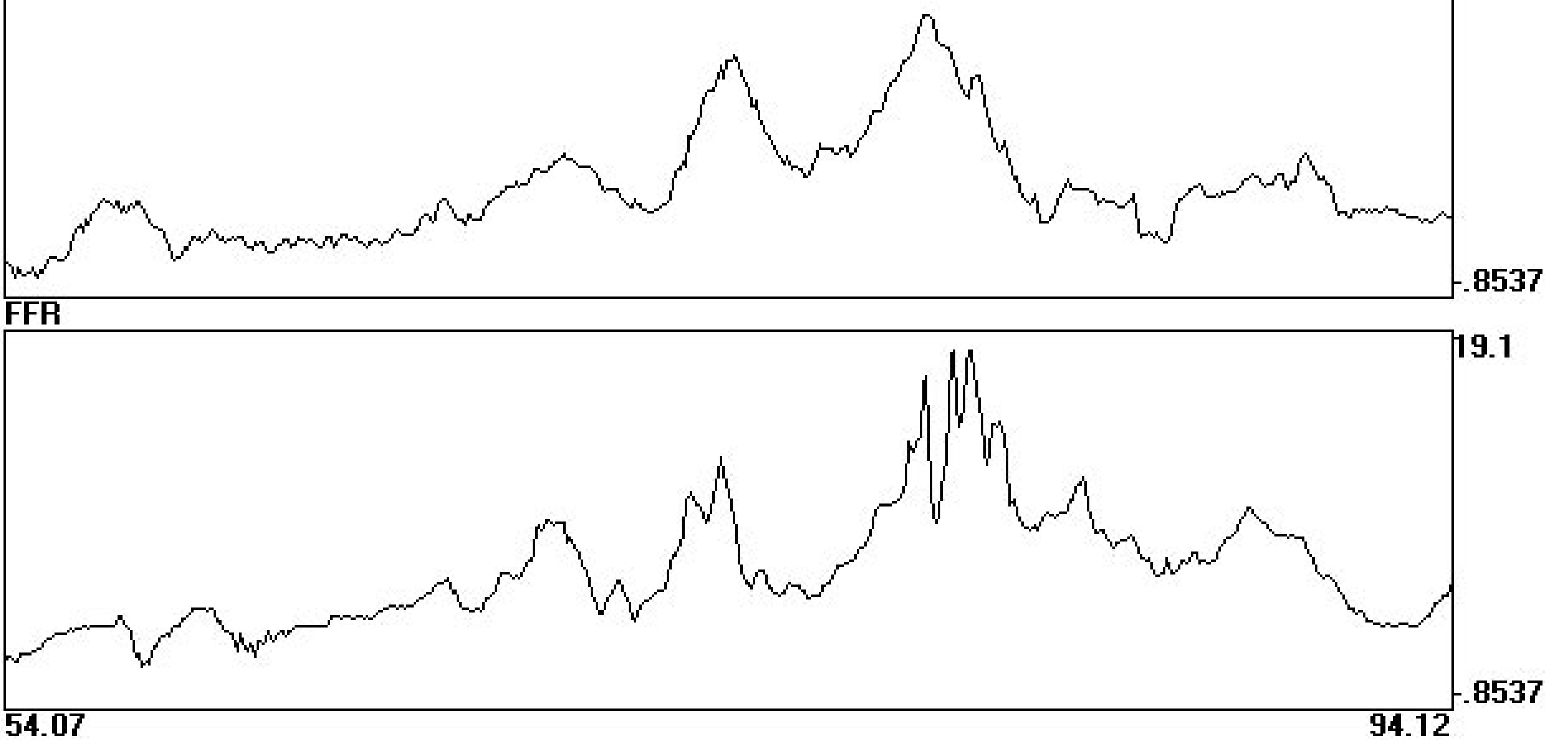

Figure 1

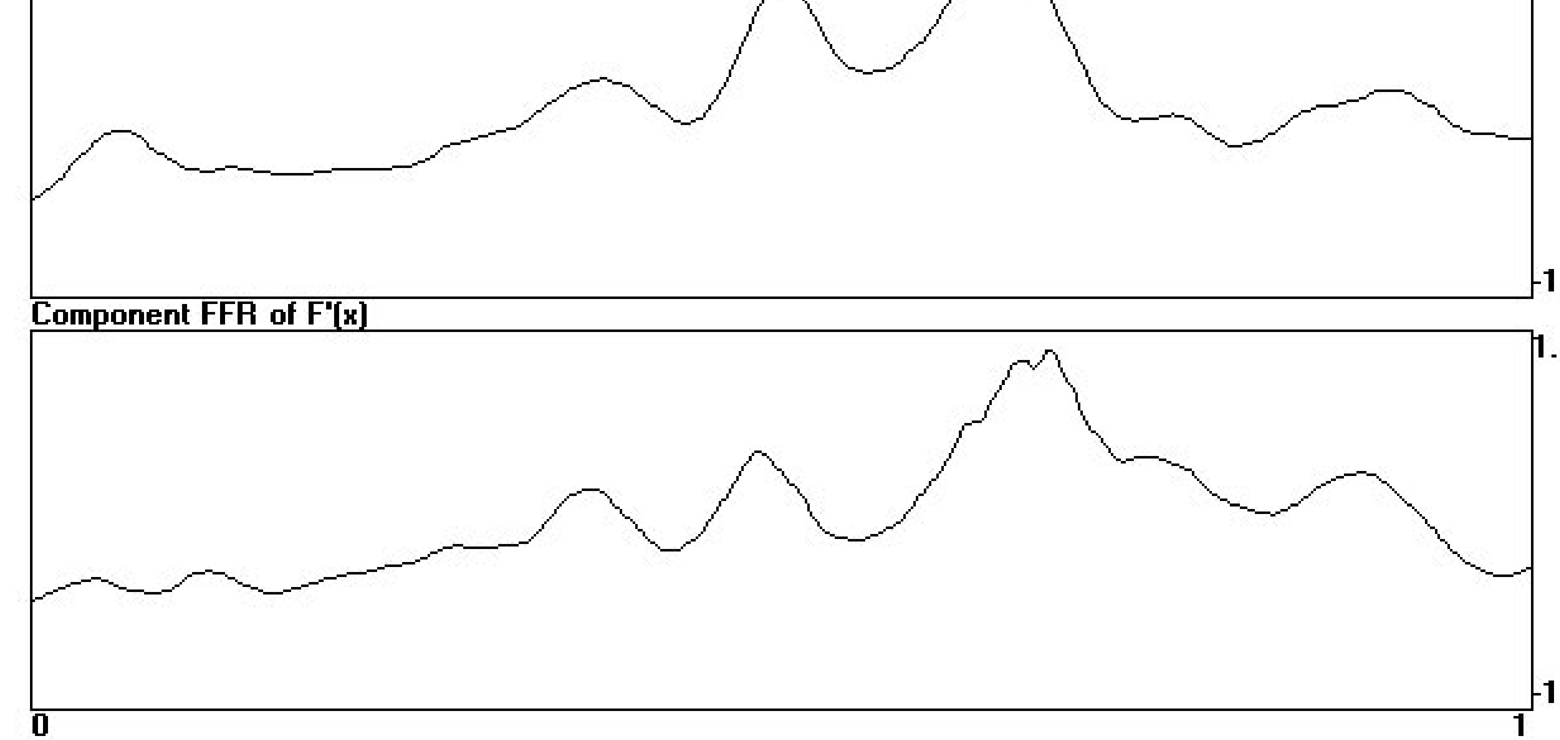

Figure 3: Rescaled components of F'(x)